

Chaperone assisted translocation

Tobias Ambjörnsson and Ralf Metzler

NORDITA (Nordic Institute for Theoretical Physics),
Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark.

E-mail: ambjorn@nordita.dk, metz@nordita.dk

Abstract.

We investigate the translocation of a stiff polymer through a nanopore in a membrane, in the presence of binding particles (chaperones) that bind reversibly to the polymer on both sides of the membrane. A bound chaperone covers one (univalent binding) or many (multivalent binding) binding sites. Assuming that the diffusion of the chaperones is fast compared to the rate of translocation we describe the process by a one-dimensional master equation. We expand previous models by a detailed study of the effective force in the master equation, which is obtained by the appropriate statistical mechanical average over the chaperone states. The dependence of the force on the degree of valency (the number of binding sites occupied by a chaperone) is studied in detail. We obtain finite size corrections (to the thermodynamical expression for the force), which, for univalent binding, can be expressed analytically. We finally investigate the mean velocity for translocation as a function of chaperone binding strength and size. For both univalent and multivalent binding simple results are obtained for the case of a sufficiently long translocating polymer.

PACS numbers: 02.50.-r, 82.37.-j, 87.16.-b

Published in Physical Biology 1, 77 (2004).

1. Introduction

The problem of polymer translocation, i.e., the transport of an oligomer or polymer (e.g., DNA, RNA or proteins) through a nanopore in a membrane, is a process of fundamental importance in biology and biotechnology. Relevant biological examples of this type of process include: translocation of proteins through the endoplasmatic reticulum, translocation of RNA through the nucleus pore membrane, the viral injection of DNA into a host and DNA plasmid transport from cell to cell through cell walls [1]. Additionally, biotechnological applications connected to membrane pore passaging, such as rapid reading of DNA base sequences [2, 3], analyte detection [4], and nanosensor applications, have been suggested. In medicine, controlled drug delivery is an ultimate goal, a crucial element of which is the passage of cell and/or nuclei membranes.

On the theoretical side there exist a number of investigations [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18], whose common approach to the translocation problem is to employ a one-dimensional description of the process using the penetration length into the pore as a slow variable (“reaction” coordinate); each translocation step is assumed to be sufficiently slow so that the polymer has time to relax to local equilibrium during the step (the instantaneous relaxation approximation [19]). The dynamics is then Markovian and can be described by a one-dimensional Smoluchowski (Fokker-Planck) or master equation [11] in terms of the slow variable (however, as pointed out in [9], this approach breaks down for very long polymers[‡]). The force appearing in the Smoluchowski equation in general has entropic (chain confinement in the pore reduces accessible degrees of freedom [12]) as well as external (electric field, chaperone binding etc.) contributions. Different theoretical studies have focused on different experimentally measurable entities: The mean translocation time is the most studied quantity [5, 9]. More detailed studies investigated the probability density of translocation times [6, 11]. Also the flux (number of polymers passing through the pore per unit time) has been theoretically investigated [12]. As pointed out already there exist certain scenarios according to which the translocation dynamics becomes subdiffusive [9, 21]. However in the present work we concentrate on a system whose dynamics is Markovian.

Two important driving forces for translocation, both in vivo and in experimental assays, are (i) an electric field across the membrane and (ii) binding particles (chaperones). In this study the focus is on the latter mechanism which appears to be particularly common for protein translocation [10, 13, 14, 22, 23, 24], but also of relevance for DNA transport through membranes [25, 26]. In the careful investigation by Simon et al. (see [22]) it was suggested that the translocation of proteins is a simple thermal ratchet process, i.e., the role of the chaperones is simply to prevent the backward diffusion through the pore, thereby speeding up translocation. In other studies the effect of the chaperones is modelled by using an effective force originating from the chemical potential difference due to the chaperones on the two sides of the

[‡] For very long polymers, the process becomes subdiffusive, and the Fokker-Planck equation may be replaced by its fractional analogue [20].

membrane [5, 6, 7, 8, 17]. More recently [10], the coupled translocation-chaperone dynamics was investigated by Brownian molecular dynamics simulation, initiating more detailed studies of the chaperone-assisted translocation process. Although the studies above provide insights into the role of the chaperones in the translocation process, there is still no unified understanding.

In the present work we perform a detailed theoretical investigation of the chaperone assisted translocation. The main contributions are (i) our result for the finite size correction to the force; (ii) that we include the possible occurrence of chaperones on both sides of the pore; and (iii) that we consider also chaperones which are larger than the size of a binding site. The paper is organized as follows: In section 2 we estimate the different relevant timescales of the problem, and in particular we distinguish between reversible and irreversible binding to the polymer. We also provide a general framework in terms of a master equation, which allows a theoretical description of the translocation dynamics. In section 3 we calculate the force on a stiff polymer in the reversible binding regime by the appropriate statistical mechanical average over the chaperone states. The polymer is divided into M equidistant binding sites to which the chaperones can bind, and we distinguish between the cases when a bound chaperone covers one (univalent binding) or several binding sites (multivalent binding). For the case of univalent binding we recover previous thermodynamical results for the force, however with a finite size correction. In section 4 we use the results from the previous section in order to compute the mean velocity of the polymer through the pore. For the case of sufficiently long stiff polymers simple results are obtained for both univalent and multivalent binding. In section 5 we compare the effectiveness of the chaperone assisted translocation to electric field driven translocation. Finally, in section 6 we give a summary and outlook.

2. General framework and relevant timescales

In this section we provide a general framework for describing sufficiently slow translocation dynamics in terms of a one-dimensional master equation. By estimating relevant time-scales we distinguish between three translocation regimes.

The geometry considered in this study is shown in figure 1: A rod-like polymer is translocating through a narrow pore in a membrane (the pore is usually a few nanometer in size, corresponding to 10-15 nucleotides in the case of ssDNA and RNA translocation through the α -hemolysin channel). On each side of the membrane there are chaperones which can bind to the polymer. We here use the binding site size σ (see figure 1) as the basic unit of length; for the case of polynucleotides σ may be the size of a base, and for unfolded proteins σ may correspond to the size of an aminoacid. The translocation process is then described by the variable m , which is the number of binding sites on side B (the distance the polymer has entered into side B is $x = m\sigma$). The total number of polymer binding sites is $M = L/\sigma$ (L is the length of the polymer) and therefore for a given m the number of binding sites on side A is $M - m$. Denote by $P(m, t)$ the probability that the polymer has passed with m binding sites into side B at time t . We

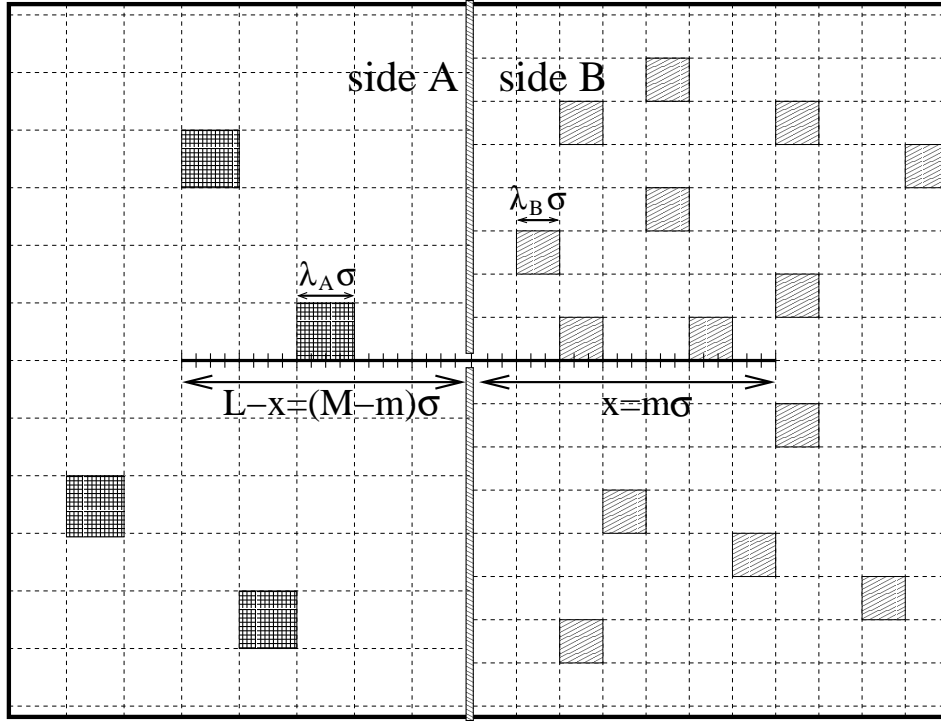


Figure 1. Translocation geometry in the presence of chaperones: A stiff polymer of length L is translocating through a pore in a membrane. The filled boxes of the two sides are chaperones. Each chaperone on side A (side B) occupies a volume v_{0A} (v_{0B}), and binds to the polymer with a binding energy ϵ_A (ϵ_B). The volume of the compartment of side A (side B) is V_A (V_B). The size of a binding site is σ and when a chaperone is attached to the polymer on side A (side B) it occupies λ_A (λ_B) binding sites. The total number of available binding sites is $M = L/\sigma$. The number of binding sites on side B is $m = x/\sigma$, and the number of binding sites on side A is $M - m = (L - x)/\sigma$.

assume that $P(m, t)$ satisfies a master equation: §

$$\begin{aligned} \frac{\partial}{\partial t} P(m, t) = & t^+(m-1)P(m-1, t) + t^-(m+1)P(m+1, t) \\ & - (t^+(m) + t^-(m)) P(m, t). \end{aligned} \quad (1)$$

Equation (1) is our starting point for studying the chaperone assisted translocation process. Transitions can occur only one step in the forward or backward direction (i.e, m is increased or decreased by one)¶. The transition probability for forward and backward motion is described by the transfer coefficients $t^+(m)$ and $t^-(m)$, respectively. In order to have a complete description we must specify $t^+(m)$ and $t^-(m)$ in terms of

§ A general master equation has the form [27] $\partial P(m, t)/\partial t = \sum_{m'} (W(m|m')P(m', t) - W(m'|m)P(m, t))$, where $W(m|m')$ are the transition probabilities per unit time. Assuming that transitions can only occur in unit steps, i.e., $W(m|m') = t^+(m')\delta_{m, m'+1} + t^-(m')\delta_{m, m'-1}$ we obtain equation (1).

¶ This assumption is reasonable in view of the fact that the nanopore only allows a 1D array of monomers (aminoacids, nucleotides) and that the passage is associated with a friction.

the fundamental parameters of the problem. We choose these entities such that the Smoluchowski (Fokker-Planck) equation is recovered in the limit $\sigma \rightarrow 0$. This requires ¶

$$\begin{aligned} t^+(m) &= \frac{1}{\tau_0} \left(1 + \frac{F(m)}{2F_0} \right), \\ t^-(m) &= \frac{1}{\tau_0} \left(1 - \frac{F(m)}{2F_0} \right), \end{aligned} \quad (2)$$

with a characteristic time $\tau_0 = \sigma^2/D$ and characteristic force $F_0 = k_B T/\sigma$, where D is the diffusion constant for the polymer and T is the temperature of the solvent (k_B is the Boltzmann constant). $F(m)$ is the force acting on the polymer; in this work the focus is on obtaining relevant expressions for $F(m)$ when both sides of the membrane contain a certain population of binding particles (chaperones), which can bind reversibly to the polymer, as is illustrated in figure 1. The chaperones are larger than the pore size, so that there is no exchange of chaperones between the two membrane sides. Typically the translocation is driven by the binding of chaperones on the exit side (side B). However, in vivo, chaperones are often present on both membrane sides (possibly the role of the chaperones on the entrance side is to unfold proteins before translocation [22]), and we therefore allow for the presence of chaperones in both compartments in figure 1. For instance, protein import into mitochondria requires the presence of chaperones both on the cytosolic side (proteins belonging to the cytosolic hsp70 family) and on the mitochondrial side (mitochondrial hsp70) [1]. The use of a time-independent force in equation (2) relies on the assumption that the chaperone dynamics is fast compared to a characteristic translocation timescale (see below). For a *flexible* polymer (not considered in detail in this study) a time-independent force requires in addition that the relaxation time of the polymer is small compared to relevant translocation times.

Let us now investigate how fast a polymer translocates through the pore, assuming, as in previous approaches, that once in the pore the polymer is not allowed to fully retract to the entrance side. Keeping in mind that the coordinate m is then confined to the interval $0 \leq m \leq M$, and that we have a reflecting barrier at $m = 0$, and an absorbing state at $m = M + 1$, the mean translocation time, for a process described by the master equation (1), becomes [27]

$$\tau = \tau_0 \sum_{m=0}^M \left(\Phi(m) \sum_{m'=0}^m \frac{1}{t^+(m')\Phi(m')} \right), \quad (3)$$

with

$$\Phi(m) = \prod_{u=1}^m \frac{t^-(u)}{t^+(u)}. \quad (4)$$

¶ In the continuum limit $\sigma \rightarrow 0$, equations (1) and (2) satisfy the following Smoluchowski (Fokker-Planck) equation [27] $\partial P(x, t)/\partial t = D\partial/\partial x \{-F(x)P(x, t)/k_B T + \partial P(x, t)/\partial x\}$ where $x = m\sigma$ is the distance the polymer has entered into side B. In the equation above the Einstein relation $D = k_B T/\xi$ (ξ is the friction constant for the polymer) is implicit.

We have above assumed that at time $t = 0$ we start with an initial condition at $m = 0$.⁺ For later reference we introduce an average velocity

$$\langle v \rangle \equiv \frac{L}{\tau} = v_0 \frac{\tau_0}{\tau} M, \quad (5)$$

where we have defined a characteristic velocity $v_0 \equiv D/\sigma$. It is often more illustrative to use $\langle v \rangle$ when discussing the results rather than the mean translocation time τ . For the case of a constant force $F(m) = F$ (so that the transition probabilities are independent of m , $t^+(m) = t^+$ and $t^-(m) = t^-$) the mean translocation time can straightforwardly be calculated using equations (2) and (3), see Appendix A. The corresponding mean velocity, for large M (see equation (A.2)), becomes

$$\langle v \rangle / v_0 = t^+ - t^- = F/F_0. \quad (6)$$

Thus, for a constant force the mean velocity of a long polymer through the pore is simply proportional to the force acting on the polymer, i.e., the average motion is equivalent to classical motion, as it should. As we will see in the subsequent section the force due to the chaperones is such that it becomes constant for $m > m_0$, where m_0 is some characteristic finite size correction length. For a sufficiently long polymer the finite size correction is negligible and the expression above can be used to obtain the mean velocity. However, very frequently the translocating chains are relatively short (for instance, small proteins are ~ 60 aminoacids long, see reference [1] pp. 117-118; or in vitro studies of single-stranded DNA translocation concern chain lengths well below 100 bases [3]), and the finite size corrections do come into play.

There are three relevant time scales associated with the problem: The time τ_{diff} needed for the polymer to diffuse a distance of the order of the binding site length σ ; the typical time τ_{unocc} a binding site stays unoccupied; and the characteristic time τ_{occ} that a binding site remains occupied. Let us estimate these different time scales, assuming that there are no chaperones on the entrance side (side A) for simplicity.* The time needed for the translocating polymer to diffuse a distance σ is simply $\tau_{\text{diff}} = \tau_0/2 = \sigma^2/2D$. Taking $\sigma \simeq 1$ nm and $D \simeq 0.1$ nm²/s[‡] we find $\tau_{\text{diff}} \simeq 5$ s. We now consider τ_{unocc} and τ_{occ} . Denote by D_c the bulk diffusion constants for the chaperones and by c_0 the bulk concentration of chaperones. The chaperones bind to the polymer site with a binding energy ϵ (< 0). Clearly the probability that a binding site is occupied depends on both the concentration and the binding energy, and we will see in the next section (see also Appendix B) that we can form a dimensionless number $\kappa = c_0 K^{\text{eq}}$ which is a

⁺ In the continuum limit equation (3) becomes [27] $\tau \simeq \int_0^L dx'' \exp(\beta G(x'')) \int_0^{x''} dx' \exp(-\beta G(x')) / D$ where L is the (contour) length of the polymer, $G(x) = -\int_0^x F(x') dx'$ and $x = m\sigma$.

* We do not consider the characteristic time associated with one dimensional diffusion of binding proteins *along* the polymer. For instance, for non-specific binding of DNA transcription factors this introduces yet another time scale [28].

[‡] We here use the estimated effective diffusion constant $D \simeq 0.1$ nm²/s for protein import into mitochondria in reference [29], obtained by comparison to experimental translocation times. Notice that this value for the diffusion constant is orders of magnitude smaller than that of a freely diffusing polymer [29], probably due to polymer interactions with the pore.

relevant measure of the binding strength, where we have defined an equilibrium binding constant $K^{\text{eq}} = v_0 \exp(\beta|\epsilon|)$ (v_0 is the typical size of the chaperones, $\beta = 1/(k_B T)$, with k_B being the Boltzmann constant and T the temperature of the solvent as before). Considering univalent binding for simplicity, the equilibrium probability that a binding site is occupied is (see Appendix C) $P_{\text{occ}}^{\text{eq}} = \kappa/(1 + \kappa)$, and the probability that a binding site is unoccupied is therefore $P_{\text{unocc}}^{\text{eq}} = 1 - P_{\text{occ}}^{\text{eq}} = 1/(1 + \kappa)$. At equilibrium the ratio of these probabilities is equal to the ratio between the times τ_{occ} and τ_{unocc} , i.e. we have $\tau_{\text{occ}}/\tau_{\text{unocc}} = \kappa$. We now proceed by obtaining τ_{unocc} , which then through the above relation also determines τ_{occ} : Consider a binding site, which initially is vacant. If we assume that κ is not too small, then as soon as one chaperone is at the binding site it becomes trapped and the binding site occupied. The distance between chaperones in solution is $R \sim c_0^{-1/3}$. It suffices for a chaperone to diffuse a distance of the order R for any one chaperone to attach to the binding site (provided κ is sufficiently large), this takes a time $\tau_{\text{unocc}} \sim R^2/D_c \sim 1/(c_0^{2/3} D_c)$, which then determines the characteristic time an initially vacant binding site stays unoccupied (see reference [30] for a more thorough investigation of this problem).^{††} Taking $D_c \simeq 10^6 \text{ nm}^2/\text{s}$ and $c_0 \simeq 10 \text{ } \mu\text{M}$, we find $\tau_{\text{unocc}} \simeq 1 \text{ ms}$. Thus typically the binding time is faster than the time for the polymer to diffuse a distance of the order one binding site. We note, however, that the above estimated numerical values are very crude, and in particular that the polymer diffusion constant D may deviate substantially from the result given here (D depends on the nature of the polymer-pore interaction). We found above that $\tau_{\text{occ}} = \kappa \tau_{\text{unocc}}$, and thus if the binding strength κ is large τ_{occ} can become large, even if τ_{unocc} is small. The considerations above allow us to distinguish between three different dynamical regimes: (i) *Diffusive regime*, $\tau_{\text{diff}} \ll \tau_{\text{unocc}}, \tau_{\text{occ}}$. In this regime the diffusion through the pore is so fast that the chaperones do not have time to bind to the translocating polymer. The force in equation (2) is then essentially zero, and the mean translocation time equation (3) becomes simply $\tau = \tau_0 M^2/2$. This regime can always be reached by lowering the concentration of chaperones sufficiently. The diffusive regime corresponds to cases previously discussed (see for instance [7, 5]) and will therefore not be considered further in this investigation.

(ii) *Irreversible binding regime*, $\tau_{\text{unocc}} \ll \tau_{\text{diff}} \ll \tau_{\text{occ}}$. This regime corresponds to a situation when the particles have sufficient time to bind, however do not unbind from the polymer during the translocation. In this so called Brownian ratchet regime [22, 10, 5] the particles, at sufficiently high concentrations and binding energies, bind immediately as soon as the polymer has diffused a distance equal to the size of a chaperone (or for univalent binding, a distance equal to the distance between binding sites). The binding of chaperones prohibits backward diffusion through the pore, and the polymer can only “jump” in the forward direction. This regime can formally be obtained by letting $F = 2F_0$ in equation (2); then the backward transition probability is zero, $t^- = 0$, the

^{††}Clearly, the depletion of chaperones due to binding to a neighbouring binding site could affect the time for binding. The concentrations of chaperones is expected to be sufficiently high so that this effect will not be dominant for estimating the relevant binding time.

forward transition probability per unit time is $t^+ = 2/\tau_0$, and hence the master equation (1) becomes

$$\frac{\partial}{\partial t}P(m, t) = \frac{2}{\tau_0} (P(m-1, t) - P(m, t)). \quad (7)$$

This equation gives a coarse-grained description of the ratchet process, where only forward jumps are effectively allowed. The same type of equation appears also in the theoretical description of shot noise [27], and in the continuum limit corresponds to the forward mode of the wave equation [31]. The mean translocation time for ratchet motion becomes $\tau = \tau_0 M/2$ and hence the mean velocity (see equation (5)) is $\langle v \rangle = 2v_0$, which agrees with the result of other studies [10, 22]. The ratchet mechanism gives a decrease of the translocation time by a factor $\sigma/L = 1/M$ compared to the translocation time in the diffusive regime described in (i) above.

(iii) *Reversible binding regime*, $\tau_{\text{unocc}}, \tau_{\text{occ}} \ll \tau_{\text{diff}}$. In this regime the particles have time to bind and unbind many times during the time it takes for the polymer to diffuse a distance σ . The polymer thus has time to reach local equilibrium, and as we will see we can then obtain the force $F(m)$ by the appropriate statistical mechanics average of the chaperone states. We will in the rest of this paper investigate this case of *reversible* (both univalent and multivalent) binding more closely.

3. Force $F(m)$ in the reversible binding regime

In this section we investigate the force $F(m)$ for reversible binding of chaperones to the translocating polymer. The chaperones are assumed to cover one binding site (univalent binding) or many binding sites (multivalent binding), when attached to the polymer.

3.1. General expression for the force $F(m)$

We start by deriving a general expression for the force $F(m)$ on the translocating polymer, arising from the interaction with chaperones on the two sides of the membrane, in the reversible binding regime.

Let us obtain a statistical mechanical expression for the force on the translocating polymer. Denote by $Z(m, n_A, n_B)$ the Boltzmann-weighted number of configurations for a state specified by m , n_A and n_B , where n_A (n_B) is the number of attached chaperones on side A (side B). For two unconnected compartments (see figure 1) this statistical weight can be written as the product of the statistical weight on side A and B respectively, i.e., $Z(m, n_A, n_B) = Z_A(m, n_A)Z_B(m, n_B)$. This is a natural decomposition, as the binding proteins cannot cross the nanopore (at least not in the presence of the translocating polymer). The force then decomposes in the form

$$F(m) = F_A(m) + F_B(m), \quad (8)$$

i.e., the total force has independent contributions from side A and side B respectively. Let us proceed by writing down the appropriate statistical mechanical expression for

the force originating from side γ ($\gamma=A$ or B). We have (see reference [10])

$$\begin{aligned} \frac{F_\gamma(m)}{F_0} &= \sum_{n_\gamma=0}^{n_\gamma^{\max}} \frac{\beta \partial \ln Z_\gamma(m, n_\gamma)}{\partial m} \rho_\gamma^{\text{eq}}(m, n_\gamma) \\ &= \frac{1}{Z_\gamma(m)} \sum_{n_\gamma=0}^{n_\gamma^{\max}} \frac{\partial}{\partial m} Z_\gamma(m, n_\gamma), \end{aligned} \quad (9)$$

where $\rho_\gamma^{\text{eq}}(m, n_\gamma)$ is the equilibrium probability density for a state specified by m and n_γ . We have above used the fact that the explicit expression for the equilibrium distribution is $\rho_\gamma^{\text{eq}}(m, n_\gamma) = Z_\gamma(m, n_\gamma)/Z_\gamma(m)$, where the partition function $Z_\gamma(m)$ for side γ is obtained by summing the statistical weight over all allowed values of n_γ :

$$Z_\gamma(m) = \sum_{n_\gamma=0}^{n_\gamma^{\max}} Z_\gamma(m, n_\gamma), \quad (10)$$

where n_γ^{\max} is the maximum number of attached binding particles on side γ . Notice that this quantity depends on m . As before, we take $F_0 = k_B T / \sigma$. The force, equation (9), is given by weighting the derivatives, $-\partial \ln Z_\gamma(m, n_\gamma) / \partial m$, of the free energy of state m, n_γ by the equilibrium probability density (compare to equation (C.1)). We note here that in order for the derivative in the force expression to be well-defined, the equation for the free energy must be analytically continuable to non-integer m (the expression for the free energy considered here are expressible in terms of factorials, which can be analytically continued through Γ -functions, see next subsection). We point out that the force $F_\gamma(m)$ may in general incorporate other effects, for instance, as caused by electric fields or protein folding (see discussions in sections 5 and 6). If the polymer is flexible chain entropic effects give an additional contribution to the force. Provided that the chaperone binding is independent of the curvature of the polymer, the binding force (as calculated in the next section) and the entropic force are additive, and the entropic force expression given in reference [5] can be used. For the sake of clarity, we here neglect entropic effects, i.e., we assume a rod-like polymer. This is not a strong restriction of the model, as in the presence of a drift as exerted by the chaperones (see below), the entropic effect is expected to be negligible for most chain lengths relevant to proteins, compare reference [6].

Let us finally rewrite the force in a form convenient for obtaining $F(m)$ in the thermodynamic limit ($m \rightarrow \infty$ for side B and $M - m \rightarrow \infty$ for side A). We write

$$F_\gamma(m) = \bar{F}_\gamma(m) - \Delta_\gamma(m), \quad (11)$$

with

$$\frac{\bar{F}_\gamma(m)}{F_0} \equiv \frac{\partial Z_\gamma(m) / \partial m}{Z_\gamma(m)} = \frac{\partial}{\partial m} \ln Z_\gamma(m), \quad (12)$$

and

$$\frac{\Delta_\gamma(m)}{F_0} \equiv \frac{\partial Z_\gamma(m) / \partial m - \sum_{n_\gamma=0}^{n_\gamma^{\max}} \partial Z_\gamma(m, n_\gamma) / \partial m}{Z_\gamma(m)}, \quad (13)$$

where the partition function $Z_\gamma(m)$ is given in equation (10). The force (11) is composed of two terms. The first term, explicitly given by equation (12), is the thermodynamic expression for the force obtained by taking the derivative of the logarithm of the partition function with respect to m . As we will see in subsections 3.3 and 3.4 this term is in general independent of m , and is proportional to the chemical potential difference across the membrane. The second term, given in equation (13), is a correction term to this thermodynamic result. We note that the presence of this correction term is due to the fact that the upper limit n_γ^{\max} in the sum in (13) depends on m (if this were not the case we could move the derivative in front of the sum and $\Delta_\gamma(m)$ would be identically zero). In subsection 3.3 we show that $\Delta_\gamma(m)$ vanishes for large m for the case of univalent binding. We will therefore henceforth call $\Delta_\gamma(m)$ a finite size correction term. We note here that in order to calculate the general force expression equation (9), we must evaluate (complicated) sums involving the statistical weights $Z_\gamma(m, n_\gamma)$. However, for obtaining the force equation (12) in the thermodynamic limit (i.e., large protrusion distances), a knowledge of the partition function $Z_\gamma(m)$ suffices (as we will see in subsection 3.4, $Z_\gamma(m)$ can be straightforwardly calculated using a transfer matrix approach).

3.2. Explicit expression for the force $F(m)$

In this subsection we study in more detail the forces $F_A(m)$ and $F_B(m)$ when the two compartments contain chaperones, which bind reversibly to the translocating polymer. In particular, we obtain the forces as a function of chaperone effective binding strengths and sizes.

In order to obtain the force on the polymer we must have explicit expression for the statistical weights $Z_\gamma(m, n_\gamma)$ (see equation (9); $\gamma=A$ or B). The details of the calculation of $Z_\gamma(m, n_\gamma)$ are given in Appendix B. There are two entropic effects that must be taken into account: (i) as m increases the number of available binding sites increases on side B and vice versa; and (ii) as the number n_γ of bound chaperones increases the entropy of the surrounding “gas” decreases. We neglect the reduction of volume due to the presence of the translocating polymer. We assume that the chaperones are equal in size (univalent binding) or larger than (multivalent binding) the size of a binding site, and cover an integer λ_γ (≥ 1) number of binding sites if bound to the polymer on side γ (for instance, bacterial transcription factors cover $\sim 10 - 20$ basepairs [28]). The statistical weights then become (see equation (B.5))

$$Z_\gamma(m, n_\gamma) = \Omega_\gamma^{\text{bind}}(m, n_\gamma) \kappa_\gamma^{n_\gamma}, \quad (14)$$

where $\Omega_\gamma^{\text{bind}}(m, n_\gamma)$ denotes the number of ways of arranging n_γ particles onto the m binding sites on side B or onto the $M - m$ binding sites on side A . For the case of dilute solutions the effective binding strength κ_γ appearing above can be written:

$$\kappa_\gamma = c_\gamma K_\gamma^{\text{eq}} \quad (15)$$

where c_γ is the concentration of chaperones on side γ , and K_γ^{eq} is the equilibrium binding constant, see equation (B.8). Since the effective binding strength κ_γ is proportional to

the chaperone concentration, one can experimentally vary κ_γ by changing the latter. We note that for $\lambda_\gamma \geq 2$ there exist correlations between binding sites in the sense if one binding site is occupied, then at least one of the neighbouring binding sites is occupied. In turn, a binding protein needing more than one binding site to actually bind cannot bind between already bound chaperones if their distance is less than λ_γ . Thus, the binding characteristics for large and small ($\lambda_\gamma = 1$) are different. Explicitly we have for side B [32, 33]

$$\Omega_B^{\text{bind}}(m, n_B) = \binom{m - (\lambda_B - 1)n_B}{n_B} = \frac{(m - (\lambda_B - 1)n_B)!}{n_B!(m - \lambda_B n_B)!}, \quad (16)$$

and similarly for side A with the replacement $m \rightarrow M - m$, $n_B \rightarrow n_A$ and $\lambda_B \rightarrow \lambda_A$. We note here that we have above neglected cooperative effects (i.e., the effect that the chaperones attached to the polymer may interact). Such effects are usually incorporated into the theory through a cooperativity parameter ω_γ [32, 35, 36], where no cooperativity corresponds to $\omega_\gamma = 1$. For $\omega_\gamma > 1$ (positive cooperativity) the chaperones bound to the polymer interact attractively, whereas for $\omega_\gamma < 1$ (negative cooperativity) the chaperones repel each other. The (somewhat lengthy) relevant expressions for $Z_\gamma(m, n_\gamma)$ with $\omega_\gamma \neq 1$ can be found in reference [32]. In subsection 3.4, we revisit the problem and show that the partition function $Z_\gamma(m)$ for large m may be straightforwardly obtained including cooperativity, allowing for a determination of the force $F(m)$ (see equation (12)) in the thermodynamic limit.

Equations (14), (15) and (16) for the statistical weight $Z_\gamma(m, n_\gamma)$ completely determine the effective force $F_\gamma(m)$ (see equation (9)). Combining equations (9) and (14) we straightforwardly obtain the force on the polymer from side A :

$$\frac{F_A(m)}{F_0} = - \sum_{n_A=0}^{n_A^{\max}} \rho_A^{\text{eq}}(m, n_A) \{ \Psi(M - m - (\lambda_A - 1)n_A + 1) - \Psi(M - m - \lambda_A n_A + 1) \}. \quad (17)$$

Similarly the force from side B is:

$$\frac{F_B(m)}{F_0} = \sum_{n_B=0}^{n_B^{\max}} \rho_B^{\text{eq}}(m, n_B) \{ \Psi(m - (\lambda_B - 1)n_B + 1) - \Psi(m - \lambda_B n_B + 1) \}, \quad (18)$$

where $\Psi(z) = d \ln \Gamma(z) / dz = \Gamma'(z) / \Gamma(z)$ is the Ψ -function [34], and we have analytically continued the factorials appearing in equation (16) using Γ -functions. The maximum number of particles that can attach to the polymer on side B is $n_B^{\max} = [m / \lambda_B]$, i.e., it is the largest integer smaller than or equal to m / λ_B . Similarly for side A the maximum number of attached chaperones is $n_A^{\max} = [(M - m) / \lambda_A]$. We notice that the force from side B is zero, as it should, when the chain does not protrude at that side, i.e., we have $F_B(m = 0) = 0$. Also, since $\Psi(z)$ is an increasing function with z , the force from side B is positive $F_B(m) \geq 0$, whereas the force from side A is negative $F_A(m) \leq 0$. Equations (17) and (18) are general expressions for the force, and are convenient for numerical computations. We point out that in general the force $F(m)$ depends on five dimensionless variables: the binding strengths κ_A and κ_B , the relative sizes of the chaperones λ_A and λ_B , as well as the effective length $M = L / \sigma$ of the polymer. In the

next two subsections we derive simplified approximate results for the cases: (i) univalent binding ($\lambda_\gamma = 1$), and (ii) general multivalent binding and long polymers. In the latter subsection we also revisit the problem of cooperative effects ($\omega_\gamma \neq 1$). In these two subsections we compare the results to the exact expressions for the force, equations (17) and (18).

3.3. Univalent binding

In this subsection we consider the case of univalent binding, $\lambda_\gamma = 1$, and obtain the force in the thermodynamic limit. We also derive an approximate expression for the finite size correction to the force.

For univalent binding $\lambda_\gamma = 1$ we have $n_B^{\max} = m$ and $n_A^{\max} = M - m$ and therefore $\Omega_\gamma^{\text{bind}}(m, n_\gamma) = n_\gamma^{\max}! / (n_\gamma! (n_\gamma^{\max} - n_\gamma)!)$ (see equation (16)). The partition functions $Z_A(m)$ and $Z_B(m)$ can straightforwardly be calculated using equation (14). We find

$$Z_\gamma(m) = \sum_{n_\gamma=0}^{n_\gamma^{\max}} Z_\gamma(m, n_\gamma) = (1 + \kappa_\gamma)^{n_\gamma^{\max}}, \quad (19)$$

where we have used the binomial theorem [34]. This equation is a standard result for the partition function for univalent, non-cooperative binding to a polymer [37]. Notice that the dependence on the chaperone concentration and the binding energies appear only through the quantity κ_γ (see equation (15)). Let us now calculate the force $\bar{F}_\gamma(m)$ in the thermodynamic limit. Using equations (12) and (19) we find:

$$\frac{\bar{F}_\gamma(m)}{F_0} = \frac{\bar{F}_\gamma}{F_0} = \pm \ln(1 + \kappa_\gamma), \quad (20)$$

where the plus sign corresponds to side B , and the minus sign corresponds to side A . We notice that the thermodynamic force is independent of m , and the expression above can be viewed as a chemical potential difference across the membrane, compare references [10, 17, 26].

We proceed by considering the finite size correction to the force, equation (13). Replacing the sum over n_γ in equation (13) by an integration and using Leibniz's theorem for differentiation of an integral [34] as well as the fact that $Z(m, n_\gamma^{\max}) = \kappa_\gamma^{n_\gamma^{\max}}$ we find (for side B)

$$\frac{\Delta_B(m)}{F_0} \approx \left(\frac{\partial n_B^{\max}}{\partial m} \right) \frac{Z_B(m, n_B^{\max})}{Z_B(m)} = \left(\frac{\kappa_B}{1 + \kappa_B} \right)^m = \exp(-m/m_{0B}), \quad (21)$$

where $m_{0B} = 1 / \ln([1 + \kappa_B] / \kappa_B) = 1 / \ln(f_B^{-1})$ in terms of the filling fraction f_B of the polymer on side B as contained in equation (C.2). The finite size correction for side A , $\Delta_A(m)$, is obtained in an identical fashion. The finite size correction is exponentially decreasing with increasing m , and vanishes over distances larger than $m_{0\gamma}$, where $m_{0\gamma}$ is determined by the filling fractions on the two sides; for large (small) filling fraction, i.e., large (small) κ_γ , the correction decays slowly (rapidly) with m . Figure 2 shows the above expression for the force (equations (11), (20) and (21)) together with the exact result (equation (18)). Notice that the result above, i.e. the expression the force as given by

equations (11), (20) and (21), captures the decrease of the force with decreasing m , but does not fully agree with the result from the general expression for the force (18), due to the continuum approximation leading to equation (21). We point out that a decrease of $F_B(m)$ for decreasing m agrees with the molecular dynamics simulations in reference [10]. The results obtained in this subsection show that for a long ($m > m_{0B}$) polymer the force can be calculated using the thermodynamic expression $\bar{F}_\gamma(m)/F_0 = \partial \ln Z_\gamma(m)/\partial m$, but for short polymers finite size corrections become relevant, and one has to resort to the exact expressions for the force, equations (17) and (18). In figure 2, this fact is demonstrated for the case $\lambda_B = 12$, a typical value for binding proteins.

3.4. General case, large m

In this subsection we show that for large m the force $F_\gamma(m)$ can be obtained through the solution of an algebraic equation for general λ_γ . The approach allows us to revisit the problem of cooperativity ($\omega_\gamma \neq 1$) in a straightforward manner.

In order to obtain the force for large m , a knowledge of the partition function $Z_\gamma(m)$ suffices (see equation (12)). Rather than using the combinatorial approach given in the previous subsections the partition function can more conveniently be obtained using the approach pursued in reference [35] (see also [39]): In general the partition function $Z_\gamma(m)$ can be written as [35]

$$Z_\gamma(m) = \sum_{j=1}^{\lambda_\gamma+1} \alpha_j \Lambda_j^m, \quad (22)$$

where Λ_j are the $\lambda + 1$ roots to the algebraic equation ($\gamma=A$ or B)

$$\Lambda^{\lambda_\gamma+1} - \Lambda^{\lambda_\gamma} - \omega_\gamma \kappa_\gamma \Lambda + (\omega_\gamma - 1) \kappa_\gamma = 0. \quad (23)$$

The prefactors α_j are independent of m and are explicitly given by $\alpha_j = \Lambda_j - \lambda_\gamma d\Lambda_j/d \ln \kappa_\gamma$. Equation (23) is the secular equation associated with the transfer matrix of the system [35]. Equations (22) and (23) completely determine the partition function of the system. Notice that the above approach incorporates cooperativity effects (through the cooperativity parameter ω_γ) without substantially raising the level of complexity. However, we point out that in the present approach only the partition function $Z_\gamma(m)$, and not the statistical weights $Z_\gamma(m, n_\gamma)$, can be calculated. Therefore the approach discussed in this subsection does not allow computation of the exact expression for the force (see the definition of the force, equation (9), and also equations (17) and (18)).

Let us now calculate the force for large m . Denote by Λ_γ^{\max} the largest root to the algebraic equation equation (23). Then for large m the force equation (12) becomes

$$\frac{\bar{F}_\gamma(m)}{F_0} = \frac{\bar{F}_\gamma}{F_0} \approx \pm \ln \Lambda_\gamma^{\max}, \quad (24)$$

where the plus sign corresponds to $\gamma = B$, and the minus sign to $\gamma = A$. The force in the thermodynamic limit is hence proportional to the logarithm of the largest root Λ_γ^{\max} . Equation (23) can straightforwardly be solved on a computer, and hence the

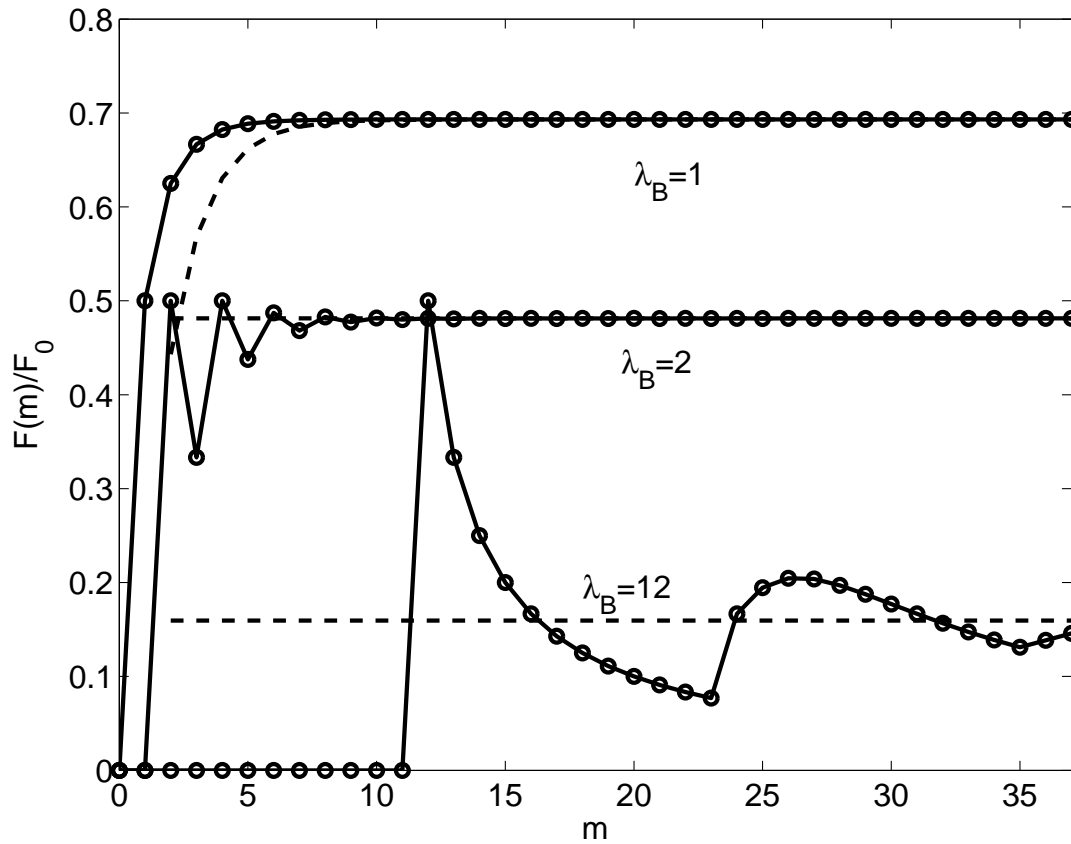


Figure 2. The force $F(m)/F_0$ in units of $F_0 = k_B T / \sigma$ (T is the temperature of the solvent, k_B is the Boltzmann constant and σ is the binding site size, see figure 1). The dashed lines correspond to approximate results: For the cases of univalent binding $\lambda_B = 1$ the upper dashed line is the analytic result as contained in equation (20). For the case $\lambda_B = 2$ we have included a plot (middle dashed line) of the approximate result contained in equation (25). For $\lambda_B = 12$, a typical value for many DNA-binding proteins, the lower dashed line corresponds to the result obtained through equations (23) and (24). The remaining curves corresponds to the results obtained through the exact expression equation (18), for: $\lambda_B = 1$ (upper curve), $\lambda_B = 2$ (middle curve) and $\lambda_B = 12$ (lower curve). The binding strength (see equation (15)) was taken to be $\kappa_B = 1$. No binding particles were assumed to be present on side A . Notice that for small m (and $\lambda_B \geq 2$) the force “oscillates” with a period λ_B . For large m the force approaches a constant value. The onset of the force is where $m = \lambda_B$, i.e., the force is zero unless there are sufficiently many binding sites on side B to accommodate at least one chaperone. The solid lines are only meant to guide the eye. For typical chaperone sizes and translocating polymer lengths the finite size corrections may become quite relevant.

determination of the force for large m is a simple matter. For univalent, non-cooperative ($\omega_\gamma = 1$) binding equation (23) becomes a first order algebraic equation which can be easily solved. The force as obtained from this solution together with equation (24) agrees with equation (20) as it should. In the previous subsection we showed that the force in the thermodynamic limit is independent of m for the case of univalent, non-cooperative binding. The result above proves that the force, for large m , is independent on m for general values of λ_γ and ω_γ . We have in figure 2 plotted the force for multivalent binding using the exact result, equation (18), as well as the above result $\bar{F}_B(m)$. The agreement is good for large m . Notice that the exact result has an “oscillatory” behaviour with a period λ_B for small m -values. We interpret these oscillations in the following way: if m is equal to an integer multiple of λ_B the polymer can, potentially (for large binding strengths), fill the polymer and hence completely restrict backward motion (perfect ‘ratcheting’). However, when m is not an integer multiple of λ_B there must be vacant spaces in between bound chaperones (for instance for $m = 5$ the maximum number of bound chaperones is 2 for divalent binding ($\lambda_B = 2$), and hence there must be at least one vacant binding site, even for large binding strengths), and the ‘ratchet’ effect is less pronounced. In reference [32] similar types of oscillations were found in the filling fraction of a polymer as a function of m , for the case of multivalent binding.

For divalent ($\lambda_\gamma = 2$), non-cooperative binding equation (23) becomes a second order algebraic equation, which can straightforwardly be analytically solved, yielding $\Lambda = [1 \pm (1 + 4\kappa_\gamma)^{1/2}]/2$. The corresponding force (for large m) becomes

$$\frac{\bar{F}_\gamma(m)}{F_0} = \frac{\bar{F}_\gamma}{F_0} \approx \pm \ln \left(\frac{1 + (1 + 4\kappa_\gamma)^{1/2}}{2} \right), \quad (25)$$

where the plus sign corresponds to $\gamma = B$, and the minus sign to $\gamma = A$. We point out that the force for divalent binding as given by equation (25) has a different functional dependence on κ_γ compared to the case of univalent binding, see equation (20).

Let us finally obtain the force for large m , univalent binding ($\lambda_\gamma = 1$) and *including* cooperativity effects (arbitrary ω_γ) using the approach above. For this case equation (23) becomes a second order algebraic equation with the roots: $\Lambda = (1 + \omega_\gamma \kappa_\gamma)/2 \pm \{(1 + \omega_\gamma \kappa)^2/4 + (\omega_\gamma - 1)\kappa_\gamma\}^{1/2}$. Hence the force, equation (24), becomes

$$\frac{\bar{F}_\gamma}{F_0} \approx \pm \ln \left\{ \frac{1 + \omega_\gamma \kappa_\gamma}{2} + \left(\left(\frac{1 + \omega_\gamma \kappa_\gamma}{2} \right)^2 + (\omega_\gamma - 1) \kappa_\gamma \right)^{1/2} \right\}, \quad (26)$$

where we have assumed that $\omega_\gamma \geq W(\kappa_\gamma) \equiv 3[2(2 + \kappa_\gamma)^{1/2}/3 - 1]/\kappa_\gamma$ so that the roots are real. We note that $W(\kappa_\gamma)$ has a maximum value $1/2$; therefore equation (26) applies whenever the cooperativity parameter satisfies $\omega_\gamma \geq 1/2$. In addition, when $\kappa_\gamma < 1/4$ we find that $W(\kappa_\gamma)$ is negative. Hence equation (26) is valid for any value of the cooperativity parameter ω_γ provided that $\kappa_\gamma \leq 1/4$. For no cooperativity $\omega_\gamma = 1$ the above result reduces to previous results (see equation (20)). Notice that the force for the case of positive cooperativity is larger than the force for negative cooperative binding, as it should.

4. Mean velocity

In this section we study the mean velocity for the polymer translocation. In particular we find a simple form for the mean velocity for long polymers.

The mean velocity is obtained on the basis of the force obtained in the previous subsections together with equations (3) and (5). For the general case (finite sized polymers) the force expressions as contained in equations (17) and (18) must be used. For sufficiently long polymers we can ignore the finite-size effect and use equation (6) together with the thermodynamic expressions for the force derived in the previous subsections. We point out that $\langle v \rangle$ then (for non-cooperative binding) depends on four dimensionless variables: $\kappa_A = c_A K_A^{\text{eq}}$, $\kappa_B = c_B K_B^{\text{eq}}$, λ_A , λ_B (c_γ is the concentration of chaperones on side γ and K_γ^{eq} is the equilibrium binding constant for the chaperones on side γ , $\gamma=A$ or B). We can therefore for sufficiently long polymers write the mean velocity according to

$$\langle v \rangle = v_0 \left(\frac{|\bar{F}_B|}{F_0} - \frac{|\bar{F}_A|}{F_0} \right) \quad (27)$$

with $v_0 = D/\sigma$ and $F_0 = k_B T/\sigma$. The relevant expressions for the forces \bar{F}_A and \bar{F}_B were given in the previous section: (i) For the general case the forces follow equation (24) and the problem is that of determining the largest root Λ^{max} to the algebraic equation (23). (ii) For *univalent* binding ($\lambda_\gamma = 1$) we use the force expressions according to equation (26). For the case of no cooperativity, $\omega_\gamma = 1$, this equation reduces to the simple result given in equation (20). (iii) For *divalent* binding ($\lambda_\gamma = 2$) and non-cooperative interactions ($\omega_\gamma = 1$) the forces are given by equation (25). When the chaperone baths on the two sides contain chaperones of identical binding strengths $\kappa_A = \kappa_B$ and sizes $\lambda_A = \lambda_B$ the mean velocity is zero, as it should. In general, however, the size of the chaperones on side A and side B may differ $\lambda_A \neq \lambda_B$, which may lead to interesting behaviour of the mean velocity as a function of κ_A and κ_B (i.e., of the chaperone concentration on the two sides). In particular we notice that the dependence on κ_γ differs between the cases of univalent and divalent binding (see equations (20) and (25)) for non-cooperative binding. The binding strength κ_γ is proportional to the concentration of chaperones on the two sides. Thus by measuring the mean velocity as a function of concentration of chaperones, it should be possible to reveal the nature of the binding on the two sides (i.e., the values of λ_γ and ω_γ). In figure 3 we have plotted $\langle v \rangle$ as a function of κ_B for different λ_B , assuming no chaperones to be present on side A for simplicity. The solid lines in figure 3 correspond to the mean velocity as calculated using equations (3), (5) and the exact expression for the force equation (18). The dotted lines corresponds to the approximate results obtained from equations (27), (20), (25) and (24). We notice that the deviation between the above approximate results and the exact result for the mean velocity is larger for larger values of κ_γ . This originates from the fact that for large κ_γ the finite size correction to the force is more pronounced (see equation (21)). For finite sized polymers the finite size correction to the force found here thus a non-negligible plays a role in the translocation dynamics. However as we

increase the length of the polymer the approximate results (dashed lines) as given above, coincide with the exact result (solid lines).

A few words on the experimental relevance of the finite size effects are in order. From figure 2 we notice that finite size effects are prominent for the *force* for m -values up to $m < (3 - 4)\lambda_B$ for the κ_B -value chosen in the figure; thus typically the larger the chaperones (larger λ_B) the more pronounced is the finite size effect for a given polymer length. We note here that it might be possible to measure the finite size effect of the force directly; for instance by attaching a bead at one end of the polymer and trapping the bead in an optical tweezer, one might directly probe the force on the polymer due to the presence of chaperones, compare to the experimental setup in reference [25]. The force is, however, not the usual experimental observable. Instead, what is usually obtained in experiments is the mean velocity (or mean translocation time). As seen in figure 3 the finite size effects are effective also for “long” polymers ($M > (3 - 4)\lambda_B$). This is due to the fact that the mean velocity, also for long polymers, contains information about the dynamics in the small m -regime.

5. Comparison to electric field induced translocation

In this section we compare the binding assisted translocation to electric field induced translocation. As we have seen in the previous sections the characteristic velocity due to binding of chaperones along a polymer is $v_0 = D/\sigma$, where D is the polymer diffusion constant and σ is the distance between binding sites. We now compare v_0 to the velocity due to the the presence of an electrostatic voltage ΔV across the membrane. If we denote the linear charge density of the polymer (charge per unit length) by ρ , the electric force on the translocating polymer is $F_{\text{elec}} = \rho\Delta V$, and hence the velocity is

$$v_{\text{elec}} = \frac{F_{\text{elec}}}{\xi} = \frac{D\rho\Delta V}{k_B T}, \quad (28)$$

where $\xi = k_B T/D$ is the friction constant for the polymer. Setting $v_0 = v_{\text{elec}}$ we find that we need a voltage across the membrane

$$\Delta V = \Delta V_t = \frac{k_B T}{\rho\sigma} \quad (29)$$

in order to get a velocity from the electric field which equals the velocity due to the chaperones. Let us estimate the voltage for a (highly) negatively charged polymer like DNA. We then take $\rho \simeq 5$ unit charges/nm, $k_B T \simeq 26$ meV (room temperature) and $\sigma \simeq 1$ nm, which gives $\Delta V_t \simeq 5$ mV. The typical (resting) potential across the eukaryotic cell membranes is $\simeq 70$ mV. Thus for charged polymers like DNA it is “preferable” to use electric fields for efficient transport. In contrast, the linear charge density of a protein is sensitive to the aminoacid sequence and the pH of the solution; at high (low) pH a protein is typically negatively (positively) charged. Therefore, *non-specific* protein transport cannot in general rely on electric field induced translocation; this may explain in parts why nature has invented the chaperone assisting machinery.

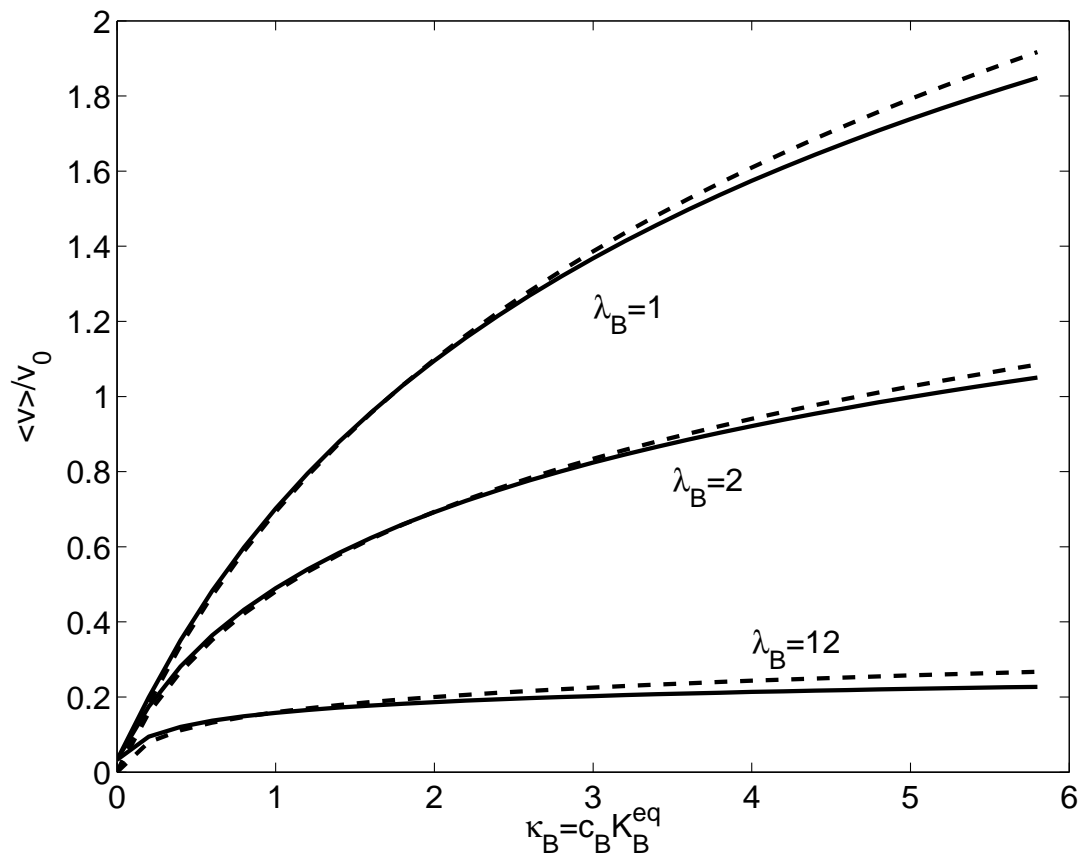


Figure 3. The mean velocity $\langle v \rangle$ for translocation of a finite sized polymer as a function of binding strength $\kappa_B = c_B K_B^{\text{eq}}$, for different relative chaperone sizes λ_B (c_B is the concentration of chaperones, and K_B^{eq} is the equilibrium binding constant). The solid lines corresponds to the mean velocities as calculated using equations (3), (5) and the exact expression for the force as contained in equation (18). The dashed curve in connection with the $\lambda_B = 1$ line, is the approximate result given in equations (27) together with equation (20). The dashed curve with the $\lambda_B = 2$ line is the approximate expression given in equation (27) and (25). For $\lambda_B = 12$ the dashed line corresponds to the result obtained through equations (23) and (24). No binding particles were present on side A. The effective length of the polymer was taken to be $M = 60$. The mean velocity increases monotonically with increasing binding strength κ_B . Notice that the deviations between the exact (solid curves) and the approximate results (dashed curves) are more pronounced for large values of κ_B .

6. Summary and outlook

We have in this work investigated the translocation of a stiff polymer through a nanopore in a membrane, in the presence of binding particles (chaperones) that bind to the polymer on both sides of the membrane. Assuming that the diffusion of chaperones is fast compared to the rate of translocation we described the process by a one-dimensional master equation. We closely investigated the translocation dynamics for the case of reversible binding to the polymer and found that the dynamics depend on whether the

chaperones bind univalently or multivalently to the polymer. For the case of univalent binding we derived an analytic finite size correction to the force exerted on the polymer by the chaperones. In general, the finite size corrections we quantified in this study may be used to extract information on the nature of the chaperones from experimental data. For long polymers a simple expression for the mean velocity of the polymer through the pore was found. We also discussed the problem of irreversible binding to the translocating polymer, as well as compared the effectiveness of binding assisted translocation to electric field driven translocation.

We want to point out that the case of perfect thermal ratchet translocation [22] (immediate irreversible binding) *cannot* be obtained by simply taking the effective binding strengths to be infinite in the results in this study. As discussed in section 2, for irreversible binding the chaperones do not have time to unbind during the translocation, rendering a thermodynamic evaluation (as we have done here) of the force inapplicable. However, we found that our master equation approach allows us to *formally* describe the case of thermal ratcheting. It will be interesting to see whether it is possible to develop a theory that covers both the reversible and irreversible binding regimes (i.e., arbitrary values of the binding strength). Possibly techniques and results from the class of the parking lot models [40, 41, 42] could prove useful.

We have in this investigation not included entropic effects due to polymer flexibility. As noted in the main text, provided that the chaperone binding does not depend on the curvature of the flexible polymer the binding force as calculated here and the entropic force are additive. For not too long flexible polymers the entropic effect could thus be included in a standard fashion (see e.g., [5]), but typically these effects in the presence of the chaperone-generated drift will be negligible for systems relevant to this study [6]. For very long polymers the dynamics changes qualitatively, and has to be modelled by a dynamical equation with memory [9]. We have also neglected the volume of the translocating polymer, which should be a fair assumption for the relevant biological systems.

We have throughout the study assumed that the binding energy for the chaperones is the same along the polymer. However proteins, RNA and DNA in general consist of heterogeneous sequences of aminoacids, bases or base-pairs respectively. It would therefore be interesting to investigate how heterogeneity in the binding energies along the polymer affects the translocation dynamics.

It has been suggested that in order for a protein to be able to translocate it has to be unfolded on the entrance side [22]. The unfolding of a protein in general requires the presence of chaperones on the entrance side; as we have seen in this study the presence of such proteins always give an opposing force compared to the translocation direction. Hence efficient translocation requires that the amount of binding proteins on the entrance side is large enough to allow unfolding, but small enough not to cause a too large opposing force. It will be interesting to see whether such an optimization concerning the concentration of “unfolders” is indeed used in nature. This situation may be improved by additional protein channels for the chaperones, by which the relative

concentration on both sides of the membrane may be actively regulated. The possibility also exists that the translocation could be driven by the refolding of the protein on the target side [22]. Alternatively, in cases where the protein is synthesised on the entrance side a built-in additional sequence can inhibit folding. This folding-preventing sequence then has to be removed on the exit side and then the folding process may assist the translocation. We note that effects similar to protein folding can occur for RNA and single stranded DNA in the form of secondary structure. In principle protein translocation could occur even for an unfolded protein, provided that the chemical or electric bias is strong enough. The translocation dynamics could in such case provide local information about the protein structure, which could find biotechnological applications, similar to RNA translocation [43].

Acknowledgments

We would like to thank John Kasianowicz and Ed Di Marzio for interesting discussions.

Appendix A. Mean translocation time for a constant force

Let us consider the mean translocation time, as given by equation (3). For the case of a constant force $F(m) = F$ the transfer coefficients, equation (2), become $t^+(m) = (1/\tau_0)(1 + F/2F_0) = t^+$ and $t^-(m) = (1/\tau_0)(1 - F/2F_0) = t^-$. Equation (3) then becomes:

$$\tau = \frac{1}{\tau^+} \sum_{y=0}^M \left(\frac{t^-}{t^+} \right)^y \sum_{z=0}^y \left(\frac{t^+}{t^-} \right)^z. \quad (\text{A.1})$$

If we now assume that $t^- < t^+$ and use the result for a geometric series (valid for $Q \neq 1$) $\sum_{z=0}^y Q^z = (1 - Q^{y+1})/(1 - Q)$ we find the following expression for the mean translocation time

$$\tau = \frac{M}{t^+ - t^-} - \frac{t^+}{(t^+ - t^-)^2} \left\{ 1 - \left(\frac{t^-}{t^+} \right)^{M+1} \right\} \rightarrow \frac{M}{t^+ - t^-} = \frac{MF_0}{F}, \quad (\text{A.2})$$

where we in the last step have assumed that $M \gg 1$. For large M (and a constant force), the mean translocation time is thus inversely proportional to the force, as it should.

Appendix B. Binding partition function

In this appendix we obtain the binding statistical weight $Z_\gamma(m, n_\gamma)$.

In order to obtain the force on the polymer we must have explicit expressions for the statistical weights $Z_\gamma(m, n_\gamma)$ (see equation (9)). This quantity is obtained by calculating the statistically averaged number of ways to attach n_γ particles on side γ , divided by a reference statistical weight Z_γ^{ref} . We choose Z_γ^{ref} as the number of states in the absence of the polymer (notice that the choice of Z_γ^{ref} is arbitrary, since it vanishes in equation (9)). Let us first calculate Z_γ^{ref} . Denote by V_A the volume of compartment A and by

V_B the volume of compartment B . Similarly, we assume a chaperone on side A (side B) to occupy a volume v_{0A} (v_{0B}). There are hence $N_A^{\text{tot}} = V_A/v_{0A}$ number of voxels to put the chaperones on side A (see figure 1), and similarly $N_B^{\text{tot}} = V_B/v_{0B}$ number of voxels to put the chaperones on side B . If we furthermore assume that there are N_A (N_B) chaperones on side A (side B), the number of ways of arranging these particles on the two sides are

$$Z_\gamma^{\text{ref}} = \binom{N_\gamma^{\text{tot}}}{N_\gamma} = \frac{(N_\gamma^{\text{tot}})!}{N_\gamma!(N_\gamma^{\text{tot}} - N_\gamma)!}, \quad (\text{B.1})$$

which then determine the reference statistical weight. Let us proceed by calculating the statistical weights $Z_\gamma(m, n_\gamma)$ in the *presence* of the polymer, neglecting the reduction of compartment volume due to the translocating polymer. As before, on side B the polymer is divided into m segments such that $m = x/\sigma$, where x is the protrusion distance on side B and σ is the size of a binding site. Similarly on side A there are $M - m = (L - x)/\sigma$ segments, where $M = L/\sigma$ is the total number of binding sites (L is the length of the polymer). We assume that the chaperones cover an integer λ_γ (≥ 1) number of binding sites if bound to the polymer on side γ . The total binding energy for an attached chaperone is denoted by ϵ_γ (< 0). The maximum number of particles that can attach to the polymer on side B is then $n_B^{\text{max}} = \lfloor m/\lambda_B \rfloor$, i.e., it is the largest integer smaller than or equal to m/λ_B . Similarly for side A the maximum number of attached chaperones is $n_A^{\text{max}} = \lfloor (M - m)/\lambda_A \rfloor$. Denote by $\Omega_B^{\text{bind}}(m, n_B)$ the number of ways of arranging n_B particles onto the m binding sites on side B ($\Omega_B^{\text{bind}}(m, n_\gamma)$ is explicitly given in the main text). To obtain the binding statistical weight for side B we have to multiply $\Omega_B^{\text{bind}}(m, n_B)$ by the Boltzmann weight associated with binding, i.e.,

$$Z_\gamma^{\text{bind}}(m, n_\gamma) = \Omega_\gamma^{\text{bind}}(m, n_\gamma) \exp(-\beta \epsilon_\gamma n_\gamma). \quad (\text{B.2})$$

In order to obtain the full statistical weight for side B , we also have to account for the fact that when n_γ number of particles are bound to the polymer there are only $N_\gamma - n_\gamma$ numbers of molecules left in the “gas” surrounding the polymer. The number of states for the “gas” (compare equation (B.1)) is

$$Z_\gamma^{\text{gas}}(n_\gamma) = \{(N_\gamma^{\text{tot}})!\} / \{(N_\gamma - n_\gamma)!(N_\gamma^{\text{tot}} - (N_\gamma - n_\gamma))!\}. \quad (\text{B.3})$$

For large N_γ (see equation 6.1.47 in reference [34]) we have the identity $(N + a)!/(N + b)! = \Gamma(N + a + 1)/\Gamma(N + b + 1) \approx N^{a-b}(1 + (a - b)(a + b + 1)/(2N) + \dots)$, where $\Gamma(z)$ is the Γ -function. Applying this result, we find that for the case when a large number of chaperones are present at the two sides:

$$Z_\gamma^{\text{gas}}(n_\gamma) = \Phi_\gamma^{n_\gamma}, \quad (\text{B.4})$$

where we have introduced the volume fraction on side γ as $\Phi_\gamma \equiv N_\gamma/(N_\gamma^{\text{tot}} - N_\gamma)$. Combining equations (B.2) and (B.4) we find that the statistical weight for side γ is:

$$Z_\gamma(m, n_\gamma) = Z_\gamma^{\text{gas}}(n_\gamma) Z_\gamma^{\text{bind}}(m, n_\gamma) = \Omega_\gamma^{\text{bind}}(m, n_\gamma) \kappa_\gamma^{n_\gamma}, \quad (\text{B.5})$$

where we have defined an effective binding strength

$$\kappa_\gamma \equiv \Phi_\gamma \exp(\beta |\epsilon_\gamma|). \quad (\text{B.6})$$

For the case of *dilute* solutions ($N_\gamma \ll N_\gamma^{\text{tot}}$) the effective binding strength can be written in the common form:

$$\kappa_\gamma = c_\gamma K_\gamma^{\text{eq}} \quad (\text{B.7})$$

where $c_\gamma = N_\gamma/V_\gamma$ is the concentration of chaperones on side γ ($\gamma=A$ or B), and

$$K_\gamma^{\text{eq}} = v_{0\gamma} \exp(\beta|\epsilon_\gamma|) \quad (\text{B.8})$$

is the equilibrium binding constant [32]. Equation (B.5) together with equation (B.7) defines the statistical weights for the two sides.

Appendix C. Filling fraction

In this appendix we consider the filling fraction of chaperones bound to the translocating polymer for the cases univalent and divalent of binding, respectively.

Since we assume that the motion of the chaperones is fast compared to the rate of translocation through the pore, the expected numbers $\langle n_A \rangle$ and $\langle n_B \rangle$ of bound chaperones on the two sides for a given m are well-defined quantities, which are simply obtained by calculating the expectation value with respect to the equilibrium distribution, i.e.,

$$\langle n_\gamma \rangle = \sum_{n_\gamma=0}^{n_\gamma^{\text{max}}} n_\gamma \rho_\gamma^{\text{eq}}(m, n_\gamma) = \frac{\partial}{\partial \ln \kappa_\gamma} (\ln Z_\gamma(m)), \quad (\text{C.1})$$

where $\gamma=A$ or B and the total expected number of bound chaperones is $\langle n \rangle = \langle n_A \rangle + \langle n_B \rangle$.

Let us calculate the expected number of bound particles on the two sides for the case of univalent $\lambda_\gamma = 1$ and non-cooperative binding $\omega_\gamma = 1$. Using equations (19) and C.1 we find $\langle n_\gamma \rangle = n_\gamma^{\text{max}} f_\gamma$ where the filling fractions are

$$f_\gamma = \frac{\kappa_\gamma}{1 + \kappa_\gamma} = 1 - (1 + \kappa_\gamma)^{-1}. \quad (\text{C.2})$$

This finding is a standard result for univalent, non-cooperative binding to a polymer [37]. We notice that $0 \leq f_\gamma \leq 1$, and that the chain becomes fully occupied, $f_\gamma \rightarrow 1$, if the binding strength is very large $\kappa_\gamma \rightarrow \infty$. When the binding strength is zero $\kappa_\gamma \rightarrow 0$ there are no chaperones bound, $f_\gamma \rightarrow 0$, as it should. For positive binding energies $\epsilon_\gamma \geq 0$ (repulsion) the filling fraction loses its meaning. Since for univalent binding (and no cooperativity) the binding sites are independent, the equilibrium probability $P_{\text{occ}}^{\text{eq}}$ that a binding site is occupied equals the filling fraction, i.e. $P_{\text{occ}}^{\text{eq}} = f_\gamma$.

We can also calculate f_γ for the case of divalent binding $\lambda_\gamma = 2$, and large protrusion distances. Using the results from subsection 3.4 we find that

$$f_\gamma \approx 1 - (1 + 4\kappa_\gamma)^{-1/2}, \quad (\text{C.3})$$

for divalent binding. We have that $0 \leq f_\gamma \leq 1$, $f \rightarrow 1$ for $\kappa_\gamma \rightarrow \infty$, and $f_\gamma \rightarrow 0$ for $\kappa_\gamma \rightarrow 0$ as it should. We notice that for divalent binding the polymer reaches its fully occupied state $f_\gamma = 1$ “slower” with κ_γ than for the case of univalent binding (see

equation (C.2)). † This is intuitively clear, as for divalent binding those configurations have to be overcome in which vacant spots of the size of one binding site have to disappear in order to reach $f_\gamma = 1$.

† Equations (C.2) and (C.3) also follows straightforwardly from the McGhee-von Hippel binding isotherm [36]:

$$\frac{f_\gamma}{\lambda_\gamma} = \kappa_\gamma (1 - f_\gamma) \left(\frac{1 - f_\gamma}{1 - (\lambda_\gamma - 1)f_\gamma/\lambda_\gamma} \right)^{\lambda_\gamma - 1}$$

which expresses the fraction f_γ of occupied binding sites as a function of κ_γ and λ_γ for binding to an infinitely long polymer. From the above expression we notice that for large chaperones $\lambda_\gamma \gg 1$ the filling fraction takes the simple form $f_\gamma = 1 - (1 + \lambda_\gamma \kappa_\gamma)^{-1}$.

- [1] Alberts B et al. 1994 *Molecular Biology of the Cell*, Garland Publishing, New York.
- [2] Henrikson S E, Misakian M, Robertson B and Kasianowicz J J 2000 Phys. Rev. Lett. **85**, 3057.
- [3] Meller A, Nivon L and Branton D 2001 Phys. Rev. Lett. **86**, 3435.
- [4] Kasianowicz J J, Henrikson S E, Weetall H H and Robertson B 2001 Anal. Chem. **73**, 2268.
- [5] Sung W and Park P J 1996 Phys. Rev. Lett. **77**, 783.
- [6] Lubensky D K and Nelson D R 1999 Biophys. J. **77**, 1824.
- [7] Muthukumar M 1999 J. Chem. Phys. **111**, 10371.
- [8] Muthukumar M 2001 Phys. Rev. Lett. **86**, 3188.
- [9] Chuang J, Kantor Y and Kardar M 2001 Phys. Rev E **65**, 011802; Kantor Y and Kardar M 2003 cond-mat/0310521.
- [10] Zandi R, Reguera D, Rudnick J and Gelbart W M 2003 Proc. Natl. Acad. Sci. USA **100**, 8649.
- [11] Flomenbom O and Klafter J 2003 Phys. Rev. E **68**, 041910.
- [12] Ambjörnsson T, Apell S P, Konkoli Z, Di Marzio E A and Kasianowicz J J 2002 J Chem. Phys. **117**, 4063.
- [13] Liebermeister W, Rapoport T A and Heinrich R 2001 J. Mol. Biol. **305**, 643.
- [14] Elston T C 2000 Biophys. J. **79**, 2235; *ibid* 2002 **82**, 1239.
- [15] Tian P and Smith G D 2003 J. Chem. Phys **119**, 11475.
- [16] Di Marzio E A and Kasianowicz J J 2003 J. Chem Phys **119**, 6378.
- [17] Kafri Y, Lubensky D K and Nelson D R 2003 cond-mat/0310455.
- [18] Meller A 2003 J. Phys. C **15**, R581.
- [19] Boehm R E 1999 Macromolecules **32**, 7645.
- [20] Metzler R and Klafter J 2000 Phys. Rep. **339**, 1.
- [21] Metzler R and Klafter J 2003 Biophys. J. **85**, 2776.
- [22] Simon S F, Peskin C S and Oster G F 1992 Proc. Natl. Acad. Sci. USA **89**, 3770.
- [23] Matlack K E S, Mothes W and Rapoport T A 1998 Cell **92**, 381; Matlack K E S, Misselwitz B, Plath K and Rapoport T A 1999 Cell **97**, 553.
- [24] Neupert W and Brunner M 2002 Nat. Rev. Mol. Cell. Bio. **3**, 555.
- [25] Salman H, Zbaida D, Rabin Y, Chatenay D and Elbaum M 2001 Proc. Natl. Acad. Sci. USA **98**, 7247.
- [26] Farkas Z, Derényi I and Vicsek T 2003 J. Phys. C **15**, 1767.
- [27] Gardiner C W 1996 *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer, Berlin.
- [28] Gerland U, Moroz J D and Hwa T 2002 Proc. Natl. Acad. Sci. USA **99**, 12015.
- [29] Chauwin J-F, Oster G and Glick S 1998 Biophys. J. **74**, 1732.
- [30] Agmon N and Gopich I V 2000 J. Chem. Phys **112**, 2863.
- [31] Landau L D, Lifshitz E M and Pitaevskii L P 1984 *Electrodynamics of Continuous Media, 2nd edition* (Oxford: Butterworth-Heinemann).
- [32] Epstein I R 1978 Biophys. Chem. **8**, 327.
- [33] McQuistan R B 1968 Nuovo Cimento **58**, 86.
- [34] Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (NIST Applied Mathematics Series, New York:Dover Publications Inc).
- [35] Di Cera E and Kong Y 1996 Biophys. Chem. **61**, 107.
- [36] McGhee J D and von Hippel P H 1974 J. Mol. Biol. **86**, 469.
- [37] Wyman J and Gill S J 1990 *Binding and Linkage - Functional chemistry of biological macromolecules*, University Science books, Mill Valley California.
- [38] A.P. Prudnikov, Y.A. Brychkov and O.I. Marichev 1986 *Integrals and series, vol. 1*, Gordon and Breach, New York.
- [39] Kong Y 2001 J. Phys. Chem. B **105**, 10111.
- [40] Krapivsky P L and Ben-Naim E 1994 J. Chem. Phys **100**, 6778.
- [41] Luthi P O, Ramsden J J and Chopard B 1997 Phys. Rev. E **55**, 3111.
- [42] Talbot J, Tarjus G and Viot P 1999 J. Phys A **32**, 2997.

- [43] Gerland U, Bundschuh R and Hwa T 2004 *Physical Biology* **1**, 19.